Basic Network Creation Games*

Noga Alon[†] Erik D. Demaine[‡] MohammadTaghi Hajiaghayi[§] Tom Leighton[¶]

Abstract

We study a natural network creation game, in which each node locally tries to minimize its local diameter or its local average distance to other nodes, by swapping one incident edge at a time. The central question is what structure the resulting equilibrium graphs have, in particular, how well they globally minimize diameter. For the local-average-distance version, we prove an upper bound of $2^{O(\sqrt{\lg n})}$, a lower bound of 3, a tight bound of exactly 2 for trees, and give evidence of a general polylogarithmic upper bound. For the local-diameter version, we prove a lower bound of $\Omega(\sqrt{n})$, and a tight upper bound of 3 for trees. The same bounds apply, up to constant factors, to the price of anarchy. Our network creation games are closely related to the previously studied unilateral network creation game. The main difference is that our model has no parameter α for the link creation cost, so our results effectively apply for all values of α without additional effort; furthermore, equilibrium can be checked in polynomial time in our model, unlike previous models. Our perspective enables simpler proofs that get at the heart of network creation games.

*A preliminary version of this paper appears in *Proceedings of the 22nd ACM Symposium on Parallelism in Algorithms and Architectures* (SPAA 2010), pp. 21–29.

†Schools of Mathematics and Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, 69978, Israel; and IAS, Princeton, NJ 08540, USA; nogaa@tau.ac.il. Supported in part by a USA Israeli BSF grant, by a grant from the Israel Science Foundation, by an ERC Advanced Grant, and by the Hermann Minkowski Minerva Center for Geometry at Tel Aviv University.

‡MIT Computer Science and Artificial Intelligence Laboratory, 32 Vassar St., Cambridge, MA 02139, USA, edemaine@mit.edu. Supported in part by NSF grant CCF-1161626 and DARPA/AFOSR grant FA9550-12-1-0423.

§Computer Science Department, University of Maryland, College Park, MD 20742; and AT&T Labs — Research, 180 Park Ave., Florham Park, NJ 07932, USA; hajiagha@cs.umd.edu. Supported in part by NSF CAREER award 1053605, NSF grant CCF-1161626, ONR YIP award N000141110662, DARPA/AFOSR grant FA9550-12-1-0423, and a University of Maryland Research and Scholarship Award (RASA).

¶Department of Mathematics, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139, USA; and Akamai Technologies; ftl@math.mit.edu

1 Introduction

In a *network creation game* (see, e.g., [FLM$^+$03, AEED$^+$06, CP05, AFM07, HM07, DHMZ12, Alb09, LPR$^+$08, DHMZ09]), several players (the nodes) collectively attempt to build an efficient network that interconnects everyone. Each player has two (selfish) goals: to minimize the cost spent building links (*creation cost*) and to minimize the average or maximum distance to all other nodes (*usage cost*). Together, these goals capture the issues of both network design and network routing.

Many different network creation games have been proposed, for example, varying which players can participate in the building of a link. All of the games, however, mediate the two objectives (creation cost and usage cost) by defining the cost of each link to be a parameter α, and minimizing the sum of creation cost and usage cost. The resulting behavior of these games seems quite intricate, and heavily dependent on the choice of α, with most bounds and proofs applying only to specific ranges of α. Furthermore, despite much effort in this area, the behavior remains poorly understood for certain ranges of α, in particular when the cost of creating a link is equivalent within a logarithmic factor to decreasing the distance to all other nodes by 1.

We introduce a basic form of the game that is at the heart of network creation games, while avoiding parameterization by a parameter α. Namely, we suppose that creation cost cannot be transformed into usage cost or vice versa, but that the cost of every edge remains equal. Equivalently, we can suppose that the uniform edge cost α is unknown. Thus, given some existing network (undirected graph), the only improving transformations that agents consider performing is the replacement of one set of edges with an equal number of other edges. We focus on the simplest form, called *basic network creation*, where each agent performs *edge swaps*: replacing an existing (incident) edge with another (incident) edge, whenever that swap improves the agent's usage cost. As a special case, the agent can swap an edge with an already existing edge, which corresponds to deleting an (extraneous) edge.

Our hope is that by simplifying the model to have no parameters, we get at the essence of the problem and shed new light on it. One motivation for this approach comes from the cache-oblivious model of computation [FLPR99, Dem02], which has successfully transformed the study of external-memory algorithms by removing parameters from the model. We believe that the techniques we introduce for our basic network creation game carry over to other network creation games (for all values of the parameter α). This opens the exciting possibility of a uniform treatment of network creation games without parameterization by α.

The study of network creation games in general, and in this paper, focuses on *equilibria*: locally stable networks in which no agent can greedily improve their situation by changing the network. In general, there are many types of equilibria depending on the types of moves allowed. The most famous is *Nash equilibria*, where each agent can change their entire strategy, while fixing all other agents' strategies. In the original network creation game [FLM$^+$03], this notion corresponds to one node deleting and/or adding any number of incident edges owned by the node. In our basic network creation game, we require *swap equilibrium*, that no edge swap decreases an agent's usage cost. Swap equilibrium is a weaker condition in that it only considers single edge swaps, instead of changing all incident edges; but it is stronger in that it allows a node to swap any incident edge, not just those that it owns.

The study of equilibria in network creation games focuses on the *price of anarchy*: the worst possible overall cost in an equilibrium network divided by the best possible overall cost in any network. This ratio gives a measure of how effectively greedy agents approximate an optimal, socially planned solution. In most existing network creation games, as well as ours, the price of anarchy turns out to be within a constant factor of the largest possible diameter of an equilibrium network, as proved in [DHMZ12]. Thus we arrive at the central question: how effectively greedy agents trying to minimize cost arrive at a low-diameter network. This question is important in its own right, beyond the application to price of anarchy, as it offers a first step toward understanding the structure of equilibria, in particular suggesting the emergence of a small-world phenomenon.

Another motivation for our basic network creation game is that Nash equilibria in network creation games are actually unrealistic: computationally bounded agents cannot even tell if they are in a Nash equilibrium (the problem is NP-complete) [FLM+03], and thus cannot tell whether they want to change their local strategy while not changing all others. In a network creation game of computationally bounded agents, it is much more reasonable to assume that an agent can only weigh a constant or sublogarithmic number of edges against each other. The same motivation led Laoutaris et al. [LPR+08] to consider a related model. Our approach is to focus on the most general extreme of these models, where agents can weigh only one edge against another edge. Questions of the form "would I rather have this edge instead of this edge?" seem natural local decisions for agents to make. In addition, the theoretical advantage of this approach is that any bounds we obtain apply equally well to all other computationally bounded models as well. When we find that diameters can actually be large ($\Omega(\sqrt{n})$) in the most general model, namely Theorem 12 in Section 4, we also consider the effect of a more powerful agent that can weigh more (up to $\Theta(\lg n/\lg\lg n)$) edges against each other, and how this improves the diameter.

Problem statement. More formally, we define two *basic network creation games*. As in previous work [DHMZ12], we consider two possible definitions of the usage cost of a node: sum and max. For the sum version, we define a graph to be in *sum equilibrium* if, for every edge vw and every node w', swapping edge vw with edge vw' does not decrease the total sum of distances from v to all other nodes.

For the max version, we define the *local diameter* of a vertex v to be the maximum distance between v and any other vertex. (A similar model was also considered in [LPR+08].) We define a graph to be in *max equilibrium* if, for every edge vw and every node w', swapping edge vw with edge vw' does not decrease the local diameter of v, and furthermore, deleting edge vw strictly increases the local diameter of v. The latter condition is equivalent to the following graph property: a graph is *deletion-critical* if deleting any edge strictly increases the local diameter of both of its endpoints. A closely related property is the following: a graph is *insertion-stable* if inserting any edge does not decrease the local diameter of either endpoint. If a graph is both insertion-stable and deletion-critical, then it is certainly in max equilibrium. In our lower-bound constructions for the max version, we design graphs that are both insertion-stable and deletion-critical, as these properties are even stronger than max equilibria.

Note that both sum and max equilibria can be detected easily in polynomial time, even locally by each agent: simply try every possible edge swap and deletion. Thus these equilibria are more natural for computationally bounded agents.

Our results. For the sum version, we prove in Section 3.2 an upper bound of $2^{O(\sqrt{\lg n})}$ on the diameter of sum-equilibrium graphs. This result is related to a previous result for unilateral network creation games [DHMZ12]. Our result effectively gets to the essence of the previous result, using a simpler proof that is independent of α.

We conjecture that the diameter of sum equilibrium graphs is polylogarithmic, and offer interesting evidence for this conjecture in Section 5. Specifically, call a graph *distance-uniform* if all vertices have almost all vertices at the same distance d. We prove that sum equilibrium graphs induce distance-uniform graphs whose diameter is smaller by at most a factor of $O(\lg^2 n)$. We conjecture that distance-uniform graphs have polylogarithmic diameter—even getting superconstant diameter seems difficult—and prove this conjecture for Cayley graphs of Abelian groups. A proof for the general case would clearly imply our conjecture about sum equilibrium graphs. This connection shows that the structure of equilibria is closely linked to a deeper, purely graph theoretic problem of independent interest.

We prove in Section 2.1 that all trees in sum equilibrium have diameter exactly 2. In other words, the only tree in sum equilibrium is the star. By contrast, we show in Section 3.1 that general sum equilibrium graphs can have diameter 3, showing a separation between trees and general graphs. It remains open whether any sum equilibrium graphs have diameter more than 3.

For the max version, we prove in Section 4 a strong lower bound of $\Omega(\sqrt{n})$ on the diameter of max equilibrium graphs (or more precisely, insertion-stable deletion-critical graphs). We also construct graphs that are both deletion-critical and stable under k insertions, meaning that the graph is stable when the agent is permitted to change any k (incident) edges. We prove a lower bound of $\Omega(n^{1/(k+1)})$ in this case, giving a smooth trade-off between diameter and computational power. In the extreme case of $k = \Theta(\lg n / \lg \lg n)$, the lower bound becomes $\Omega(\lg n)$.

For trees in the max version, we show in Section 2.2 that, in contrast to the sum version, the diameter can be as large as 3. Conversely, we prove that no diameter larger than 3 is possible.

Overall, we propose giving further attention to basic network creation game, as it seems to capture the same essence while being easier to work with and enabling more powerful techniques.

2 Trees

We start by analyzing equilibrium trees for both the sum and max versions. In both cases, we tightly characterize the maximum possible diameter: 2 for sum and 3 for max.

2.1 Sum ⇒ Diameter 2

For the sum version, we prove that there is essentially only one equilibrium tree:

Theorem 1 *If a sum equilibrium graph in the basic network-creation game is a tree, then it has diameter at most* 2*, and thus is a star.*

Proof: Suppose for contradiction that an equilibrium tree has diameter at least 3; refer to Figure 1. Thus it has two vertices v, w at distance exactly 3, inducing a length-3 shortest path $v \to a \to b \to w$. Let s_v, s_a, s_b, s_w denote the size of the subtrees rooted at v, a, b, w, respectively, counting the roots themselves. Consider two possible swaps: (1) v replaces its edge to a with an edge to b, and (2) w replaces its edge to b with an edge to a. The first swap improves v's distance to b's and w's subtrees by 1 (the unique shortest path in the tree no longer having to pass through a), and worsens v's distance to a's subtrees by 1; thus, the swap is a net win unless $s_b + s_w \le s_a$. Similarly, the second swap is a net win unless $s_v + s_a \le s_b$. For both swaps to not be net wins, we

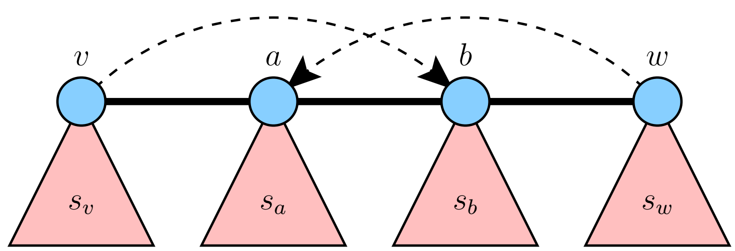

Figure 1: Illustration of Theorem 1.

must have both inequalities. Summing these inequalities, we obtain that $s_v + s_a + s_b + s_w \leq s_a + s_b$, i.e., $s_v + s_w \leq 0$, contradicting that $s_v + s_w \geq 2$ (because in particular they count v and w). $\square$

Obviously, diameter 2 can also be achieved (and is optimal), as evidenced by the star.

2.2 Max $\Rightarrow$ Diameter 3

In contrast to the sum version, max-equilibrium trees can have diameter as high as 3; see Figure 2. However, this diameter is the maximum possible. To prove this, we first need a general lemma about balance in max equilibrium graphs:

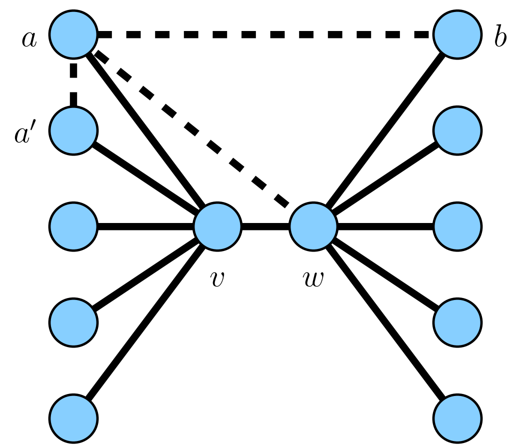

Figure 2: A tree of diameter 3 that is in max equilibrium. There are three types of edges we might try to add, shown dashed: from a leaf a to a cousin leaf a' or to a distinct leaf b or to the other root w. The only option that decreases the local diameter of either endpoint is adding aw which decreases the local diameter of a (but not w) by 1. In any swap around a, however, this addition must be combined with the deletion of edge av, which restores the original local diameter of a.

Lemma 2 *In any max-equilibrium graph, the local diameters of any two nodes differ by at most* 1*.*

Proof: Suppose vertex v has local diameter d while vertex w has local diameter at least $d+2$. Let T be a breadth-first search tree from v. We claim that w prefers to swap its edge to its parent in T with an edge to v (the root of T). Observe that this swap only decreases the depths of nodes in T, so the local diameter of v remains at most d. Thus w's local diameter decreases to at most $d+1$, because w can take a unit step to reach v and then follow v's path to any other node. This swap contradicts being in max-equilibrium. $\square$

Lemma 3 *If a max equilibrium graph has a cut vertex v, then only one connected component of $G - v$ can have a vertex of distance more than* 1 *from v.*

Proof: Let d be the local diameter of v. Let w be a vertex at distance d from v, and let W be the connected component of $G - v$ that contains w. Suppose for contradiction that there is a vertex x in $G - W$ of distance more than 1 from v. Then any path from x to W must pass through v, so it is at least 2 longer than the corresponding path from v. Therefore the local diameter of w is at least $d + 2$, contradicting Lemma 2. $\square$

Theorem 4 *If a max equilibrium graph in the basic network-creation game is a tree, then it has diameter at most* 3*.*

Proof: Suppose for contradiction that an equilibrium tree has diameter at least 4. Thus it has two vertices v, w at distance exactly 4, inducing a length-4 shortest path $v \to a \to b \to c \to w$. But then b is a cut vertex and two connected components of $G - b$ have vertices v, w of distance more than 1 from b, contradicting Lemma 3. $\square$

Therefore, there are two families of max equilibrium trees: stars (of diameter 2) and "double-stars" (of diameter 3, as in Figure 2). To be in max equilibrium, the latter type must have at least two leaves attached to each star root (v and w).

3 Sum Version

Next we analyze the case of general networks in the sum version. We start in Section 3.1 by giving the first lower bound of 3, and then turn to our sub-n^ε upper bound in Section 3.2.

3.1 Lower Bounds

Currently all examples of sum equilibrium graphs have diameter 2. Initially, Fabrikant et al. [FLM$^+$03] conjectured that sum equilibrium graphs are trees, and we have shown in Section 2 that such graphs must have diameter 2. Albers et al. [AEED$^+$06] disproved this conjecture with a cyclic sum equilibrium graph, arising from finite projective planes, but it too has diameter 2. Thus it seems reasonable to conjecture that all sum equilibrium graphs have diameter 2. Here we rule out this conjecture:

Theorem 5 *There is a diameter-*3 *sum equilibrium graph.*

First we establish a few tools for proving equilibrium in graphs of small diameter/girth. The proofs are straightforward and hence omitted.

Lemma 6 *For a vertex v of local diameter* 2*, swapping an incident edge does not improve the sum of distances from v.*

Lemma 7 *Consider a vertex v of local diameter* 3*. Adding an edge from v to a vertex w of distance r decreases the sum of the distances from v by at most $r-1$ for w and by at most* 1 *for any neighbors of w whose distance to v was* 3*.*

Lemma 8 *In any graph of girth* 4*, swapping an edge vw with edge vw' increases the distance from v to w by at least* 2*, unless w' is a neighbor of w, in which case it increases by at least* 1*.*

Proof of Theorem 5: Figure 3 illustrates the graph. One vertex a has three neighbors: b_1, b_2, b_3. Each vertex b_i has two unique neighbors other than a: $C_i = \{c_{i,1}, c_{i,2}\}$. Furthermore, for each $i \in \{1,2,3\}$, we have an additional vertex d_i connected to all of C_i. Finally, for each $i, j \in \{1,2,3\}$, $i \neq j$, we add a particular perfect matching between C_i and C_j. Between C_1 and C_2 and between C_2 and C_3, we use the obvious matching: $c_{i,1}c_{j,1}$ and $c_{i,2}c_{j,2}$. Between C_1 and C_3, we use the other matching: $c_{i,1}c_{j,2}$ and $c_{i,2}c_{j,1}$.

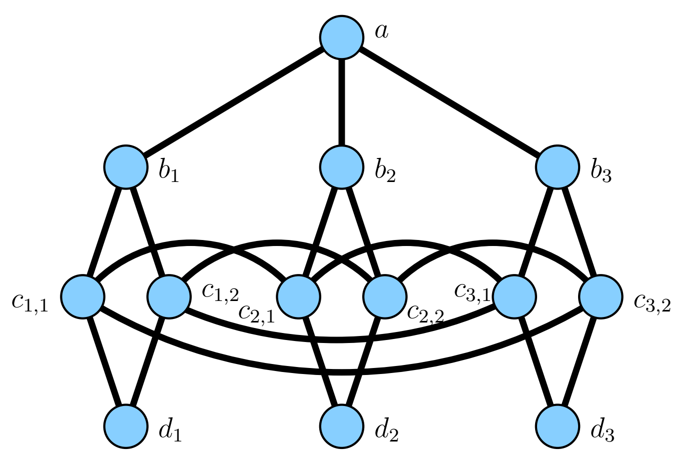

Figure 3: A diameter-3 sum equilibrium graph.

We consider all possible edge swaps around each vertex, characterizing vertices by their local diameter. By inspection, vertices a, b_i, and d_i have local diameter 3, while vertices $c_{i,k}$ have local diameter 2. In particular, the graph has diameter 3. Also observe that the graph has girth 4 (by checking that the neighbor set of each vertex is an independent set), so Lemma 8 applies.

By Lemma 6, swapping edges incident to any $c_{i,k}$ does not help. For all other vertices, we apply Lemma 7.

For vertex a, swapping an edge ab_i with $ac_{j,k}$ or ad_j decreases the sum of distances from a by at most 2: in the former case, 1 for $c_{j,k}$ and 1 for d_j; and in the latter case, 2 for d_j. Now, if $i \neq j$ or if we add edge ad_j, then by Lemma 8, the distance from a to b_i increases by at least 2, absorbing any possible benefit to the swap. If $i = j$ and we add $ac_{j,k}$, then the distances from a to b_i and to $c_{i,3-k}$ increase by 1, again absorbing the benefit.

For a vertex b_i, swapping with an edge to b_j (for $i \neq j$) or to d_i is not useful by Lemma 7, because all neighbors of b_j are already at distance at most 2 from b_i. If we swap with an edge to d_j for $j \neq i$, then by Lemma 7 we gain at most 2, but by Lemma 8 we lose at least 2, so the swap is useless. If we swap the edge $b_i a$ with $b_i c_{j,k}$ for $i \neq j$, then by Lemma 7 we gain at most 1 for $c_{j,k}$ and 1 for d_j, but by Lemma 8 we lose at least 2, so the swap is useless. If we swap an edge $b_i c_{i,k}$ with $b_i c_{j,l}$ for $i \neq j$, then by Lemma 7 we gain at most 1 for $c_{j,l}$ and 1 for d_j, but we increase the distances from b_i to $c_{i,k}$ and to at least one of its c neighbors by at least 1, so again the swap is useless.

Finally, for a vertex d_i, if we swap edge $d_i c_{i,k}$ with $d_i b_i$, then by Lemma 7 we gain at most 1 for b_i and 1 for a, but we increase the distance from d_i to $c_{i,j}$ and to each of its c neighbors by at least 1, absorbing the benefit. If we swap edge $d_i c_{i,k}$ with $d_i a$, then by Lemma 7 we gain at most 2 for a and 2 total for b_j, $j \neq i$, but by Lemma 8 we increase the distance from d_i to $c_{i,j}$ by at least 2, and we increase the distances from d_i to each of $c_{i,j}$'s c neighbors by at least 1, absorbing the benefit. If we swap edge $d_i c_{i,k}$ with anything else, then by Lemma 8 we increase the distances from d_i to $c_{i,j}$ by at least 2, and we increase the distance from d_i to at least one of $c_{i,j}$'s c neighbors by at least 1. If we swap edge $d_i c_{i,k}$ with $d_i b_j$ or $d_i d_j$ for $j \neq i$, then by Lemma 7 we gain at most 2 for b_j, and in the former case, 1 for a. If we swap edge $d_i c_{i,k}$ with $d_i c_{j,l}$ for $j \neq i$, then by Lemma 7 we gain at most 1 for $c_{j,l}$ and 1 for each of b_j and d_j. In all cases, the gain is at most the loss, so the swap is useless. $\square$

3.2 $2^{O(\sqrt{\lg n})}$ Bound

Next we prove our upper bound, which is based on the result of [DHMZ12]:

Theorem 9 *All sum equilibrium graphs have diameter* $2^{O(\sqrt{\lg n})}$.

First we need two basic results, which will find use in Section 5 as well.

Lemma 10 *Any sum equilibrium graph either has diameter at most* $2 \lg n$ *or, given any vertex* u*, there is an edge* xy *where* $d(u,x) \leq \lg n$ *and whose removal increases the sum of distances from* x *by at most* $2n(1 + \lg n)$.

Proof: Consider a breadth-first search from any vertex u in a sum equilibrium graph G. Let T denote the top $2 + \lg n$ levels of the BFS tree, from level 0 (just u) to level $1 + \lg n$. If there are any nontree edges connecting two vertices in T, then there is a cycle C whose distance from u is at most $\lg n$ and whose length is at most $1 + 2(1 + \lg n)$. In this case, each edge xy of the cycle has the property that $d(x,u) \leq \lg n$ and removing xy increases the sum of distances from x by at most $2n(1 + \lg n)$ (replacing any use of xy with the alternate path around the cycle). Thus we can assume that the graph $G[V(T)]$ induced on these top vertices is exactly the tree T.

For a vertex v in T, let T_v denote the subtree of T rooted at v. Call v *grounded* if T_v includes a node at layer $1 + \lg n$ ("the ground"). Define the *ground distance* $\mathrm{gd}(v)$ of a grounded vertex v

to be $1 + \lg n$ minus the level of vertex v, i.e., the difference in levels between v and the ground. If the root u is ungrounded, then every vertex has distance at most $\lg n$ from u, so the diameter of the graph is at most $2 \lg n$, proving the lemma. Thus we can assume that u is grounded.

We claim that every grounded vertex v other than u has $|T_v| \geq 2^{\mathrm{gd}(v)}$. In particular, applying this claim to a grounded child v of u implies that $|T| > |T_v| \geq 2^{\lg n} = n$, contradicting that the whole graph has only n vertices, and thus proving the lemma. Now we prove the claim by induction on $\mathrm{gd}(v)$. In the base case, if v is in the ground (level $1 + \lg n$), i.e., $\mathrm{gd}(v) = 0$, then $|T_v| \geq 1$ as desired because T_v includes v.

In the induction step, there are two cases. Because v is grounded, it must have at least one grounded child. If it has at least two grounded children, say a and b, then $|T_v| \geq 1 + |T_a| + |T_b| \geq 1 + 2 \cdot 2^{\mathrm{gd}(a)} = 1 + 2^{\mathrm{gd}(v)} > 2^{\mathrm{gd}(v)}$, proving the claim. Otherwise, v has exactly one grounded child, say a. Let k denote the number of ungrounded descendants of v, plus 1 to count v itself. Consider the parent p of v in T (which exists because $v \neq u$) replacing its edge pv with the edge pa. This replacement increases the distance from p to all ungrounded descendants of v, as well as v itself, by 1, increasing the sum of distances from p by (at most) k. On the other hand, ignoring these ungrounded descendants and v, the replacement shortcuts a degree-2 vertex v, so it does not increase any other distances from p. Furthermore, the replacement strictly improves the distances to all vertices in T_a by 1. Because the graph is in equilibrium, the improvement $|T_a| - k$ cannot be positive, i.e., $k \geq |T_a|$. Therefore, $|T_v| = |T_a| + k \geq 2|T_a| \geq 2 \cdot 2^{\mathrm{gd}(a)} = 2^{\mathrm{gd}(v)}$, proving the claim. □

Corollary 11 *In any sum equilibrium graph, the addition of any edge uv decreases the sum of distances from u by at most $5n \log n$.*

Proof: Suppose for contradiction that the addition of edge uv decreases the sum of distances from u by more than $5n \log n$. If the graph has diameter at most $2 \lg n$, then adding any edge can decrease each distance by at most $2 \lg n$, for a total decrease of at most $2n \lg n$. Otherwise, we find the edge xy of Lemma 10 with $d(u, x) \leq \lg n$ and whose removal increases the sum of distances from x by less than $2n(1 + \lg n)$. We claim that x prefers to replace edge xy with edge xv. The loss from deleting edge xy is at most $2n(1 + \lg n) \leq 4n \lg n$. The benefit from inserting edge xv is more than $5n \lg n - n \lg n$, because distances from u and from x differ by at most $\lg n$. The net improvement is therefore more than $5n \lg n - n \lg n - 4n \lg n = 0$, i.e., positive, contradicting that the graph is in sum equilibrium. □

Proof of Theorem 9: Consider a sum equilibrium graph G on n vertices. For any vertex u, let $S_k(u)$ denote the number of vertices at distance exactly k from u (the radius-k *sphere* centered at u). Let $B_k(u) = \sum_{i \leq k} S_k(u)$ denote the number of vertices within distance at most k from u (the radius-k *ball* centered at u). Let $B_k = \min_u B_k(u)$. We claim that

$$B_{4k} > n/2 \quad \text{or} \quad B_{4k} \geq \frac{k}{20 \lg n} B_k. \tag{1}$$

To prove (1), fix a vertex u, and assume that $B_{4k}(u) \leq n/2$. Then certainly $B_{3k}(u) \leq n/2$. Let T be a maximal set of vertices at distance exactly $3k$ from u subject to the distance between any pair of vertices in T being at least $2k + 1$. We claim that, for every vertex v of distance more than $3k$ from u, the distance of v from the set T is at most $d(u, v) - k$. Indeed, v is of distance $d(u, v) - 3k$ from some vertex at distance exactly $3k$ from u, and any such vertex is within distance $2k$ from some vertex of T, by the maximality of T. Because we assumed that at least $n/2$ vertices have distance more than $3k$ from u, by the pigeonhole principle, there are at least $n/(2|T|)$ such vertices v whose distance from the same vertex $t \in T$ is at most $d(u,v) - k$. Adding an edge from u to t improves the sum of distances from u by at least $(k-1)n/(2|T|) \geq kn/(4|T|)$. By Corollary 11, this improvement must be at most $5n \lg n$, so we conclude that $|T| \geq k/(20 \lg n)$. Now the balls of radius k centered at the vertices of T are all pairwise disjoint, all lie withing distance $4k$ of u, and each of them has at least B_k vertices (by the definition of B_k). Thus $B_{4k}(u) \geq B_k k/(20 \lg n)$, proving (1).

Now (1) easily implies that the diameter is at most $2^{O(\sqrt{\lg n})}$. First, $B_{2^{\sqrt{\lg n}}} \geq 2^{\sqrt{\lg n}}$ simply because the graph is connected. Starting from this $k = 2^{\sqrt{\lg n}}$ and applying (1), whenever we multiply k by 4, B_k increases by a factor of at least $k/(20 \lg n) \geq 2^{\sqrt{\lg n} - \lg \lg n - \lg 20} = 2^{\Omega(\sqrt{\lg n})}$, unless B_k is already more than $n/2$. Taking logarithms, $O(\sqrt{\lg n})$ such iterations suffice to reach a k where $B_k > n/2$. The diameter of the graph is then at most twice such a k, because any two vertices u, v must have overlapping balls of radius k. □

4 Max Version

Next we consider the max version, where we can prove a strong lower bound:

Theorem 12 *There is a max equilibrium graph of diameter $\Theta(\sqrt{n})$.*

Proof: Our graph G can be described roughly as 2D torus rotated 45°; refer to Figure 4. (Note, however, that a standard torus is not in max equilibrium, so the precise definition is critical.) Specifically, G has $n = 2k^2$ vertices, one for each pair (i,j) of integers where $0 \leq i,j < 2k$ and $i+j$ is even. We treat the integers as modulo $2k$; in particular, 0 and $2k$ are equivalent coordinates. Each vertex (i,j) has exactly four neighbors: $(i+1, j+1)$, $(i-1, j+1)$, $(i+1, j-1)$, and $(i-1, j-1)$. In particular, G is vertex-transitive.[1] The distance between two vertices (i,j) and (i',j') in G is exactly $\max\{d(i,i'), d(j,j')\}$, where the 1D distances are measured on the modulo-$2k$ circle: for $0 \leq i, i' \leq 2k$, $d(i,i') = \min\{|i-i'|, 2k - |i-i'|\}$. (To prove this distance formula, we simply need to observe that each coordinate can change by ± 1 in each step.)

First we show that the local diameter of every vertex is exactly k. By vertex-transitivity, it suffices to show that the local diameter of vertex (k,k) is exactly k. The

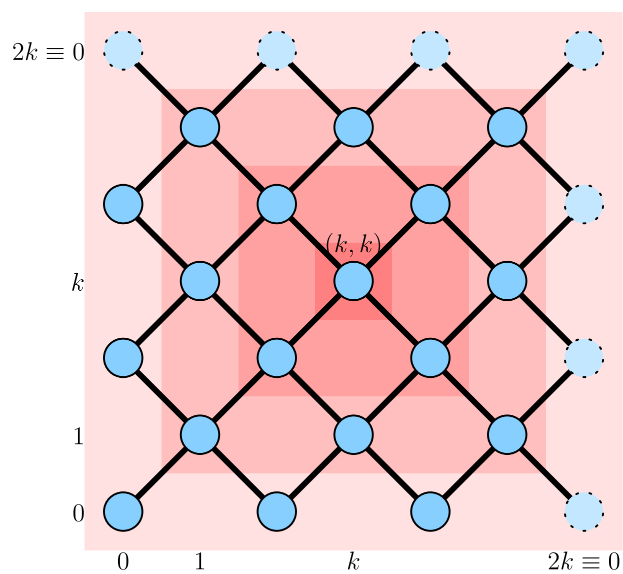

Figure 4: The $\Theta(\sqrt{n})$-diameter max equilibrium graph of Theorem 12. The rightmost and topmost columns ($x = 2k$ and $y = 2k$) are actually duplicates of the leftmost and bottommost columns ($x = 0$ and $y = 0$), respectively, and hence dotted. Shaded squares indicate distance contours from the central point (k,k).

[1]A graph is *vertex-transitive* if any vertex can be mapped to any other by a vertex automorphism, i.e., a relabeling of vertices that preserves edges.

distance between (k,k) and any vertex (i,j), where $0 \le i,j \le 2k$, is $\max\{|i-k|,|j-k|\}$, which is maximized when either i or j equals $0 \equiv 2k$.

Second we show that G is deletion-critical. By vertex-transitivity and rotational symmetry, it suffices to show that deleting the edge from (k,k) to $(k+1,k+1)$ strictly increases the local diameter of (k,k). Indeed, we claim that this deletion increases the distance from (k,k) to $(2k-1,2k-1)$ to $k+1$. Any such path must first proceed to a neighbor of (k,k), all of which have at least one coordinate of $k-1$. Thus, even in the original graph G, the distance from that neighbor to $(2k-1,2k-1)$ is k, implying that the path has length at least $k+1$.

Third we show that G is insertion-stable (and thus in equilibrium). By vertex-transitivity and rotational symmetry, it suffices to show that inserting an edge from (k,k) to $(k+i,k+j)$ does not decrease the local diameter of (k,k) for all $0 \le i,j \le k$. Consider the vertex $(2k, j-(j \bmod 2))$, which belongs to G because $2k+j-(j \bmod 2)$ is even. As we showed above, this vertex has distance k from (k,k) in the original graph G, so it remains to show that the additional edge does not make a shorter path. Such a path must first use the added edge, so it suffices to show that the distance from $(k+i,k+j)$ to $(2k, j-(j \bmod 2))$ is at least $k-1$ in the original graph G. This claim follows because the second coordinates differ by $d(k+j, j-(j \bmod 2)) = k-(j \bmod 2) \ge k-1$. □

In fact, we can also increase the dimension of the construction to d between 2 and $O(\lg n/\lg\lg n)$. Then we have a vertex for each point $(i_1, i_2, \ldots, i_d)$ where $i_1 \equiv i_2 \equiv \cdots \equiv i_d \pmod 2$, with edges to $(i_1 \pm 1, i_2 \pm 1, \ldots, i_d \pm 1)$ for all possible (independent) choices of signs for $\pm$. The resulting graph has diameter $\Theta(n^{1/d})$, which ranges from $\sqrt{n}$ to $\lg n$. This graph is deletion-critical and has the stronger property that it is stable (local diameter does not improve) under the insertion (or swapping) of up to $d-1$ edges from one vertex. The proof is similar; for insertion-stability, we can find a point that is simultaneously far from all endpoints of the added edges by devoting one coordinate to each such endpoint. This generalization is interesting because it offers a computationally tractable alternative to Nash equilibrium defined in [FLM+03], where a vertex can insert/swap an arbitrary subset of edges.

5 Connection to Distance Uniformity

Call an n-vertex graph *ε-distance-uniform* if there is a value r such that, for every vertex v, the number of vertices w at distance exactly r from v is at least $(1-\varepsilon)n$. Slightly weaker, call an n-vertex graph *ε-distance-almost-uniform* if there is a value r such that, for every vertex v, the number of vertices w at distance either r or $r+1$ from v is at least $(1-\varepsilon)n$. The following result connects high-diameter sum equilibrium graphs to high-diameter distance uniformity. We assume that the graph has more than a constant number of vertices; otherwise, the diameter is trivially constant and thus uninteresting.

Theorem 13 *Any sum equilibrium graph G with $n \ge 24$ vertices and diameter $d > 2\lg n$ induces an ε-distance-almost-uniform graph G' with n vertices and diameter $\Theta(\varepsilon d/\lg n)$ and an ε-distance-uniform graph G' with n vertices and diameter $\Theta(\varepsilon d/\lg^2 n)$.*

Proof: We use one definitional tool to establish distance uniformity. Call an ordered triple (a,b,c) of vertices *skew* if $d(a,c) > p\lg n + d(a,b)$, for a constant p to be chosen later.

First we claim that, if $p \ge 4/\alpha$, then less than an α fraction of all $n(n-1)(n-2)$ possible triples (a,b,c) are skew. For if a constant fraction $\alpha n(n-1)(n-2)$ are skew, then by averaging, there is a choice of b and c such that an α fraction of the $n-2$ choices for node a have (a,b,c) skew. Then adding edge bc improves the sum of distances from c by at least $(p\lg n - 1)\alpha(n-2) = \Omega(n\lg n)$.

By Lemma 10, and because $d > 2\lg n$, there is an edge xy where $d(c,x) \le \lg n$ and whose removal increases the sum of distances from x by at most $2n(1+\lg n)$. Because $d(c,x) \le \lg n$, adding edge xb still improves the sum of distances from x by at least $((p-1)\lg n - 1)\alpha(n-2)$. Thus, swapping edge xy with edge xb improves the sum of distances from x by at least $((p-1)\lg n - 1)\alpha(n-2) - 2n(1+\lg n)$. If $p \ge 4/\alpha$, this improvement is at least $(2-\alpha)n\lg n + 2\alpha(1+\lg n) - (2+\alpha)n - 8\lg n$. Because $\alpha \le 1$, this improvement is at least $n\lg n - 3n - 8\lg n$, which is positive for $n \ge 24$, contradicting equilibrium.

Second we claim that, if $p \ge 4/((1-2\beta)\beta)$, then for some vertex a, if we ignore the nearest βn and farthest βn nodes from a, then the remaining nodes b have distances $d(a,b)$ forming an interval of length at most $2p\lg n$. Let $[\ell_a, u_a]$ denote the interval of distances $d(a,b)$ for b among the middle $(1-2\beta)n$ nodes from a. If the claim is false, then the interval has length more than $p\lg n$ for all a. But then we claim that a constant fraction of (a,b,c) triples are skew, contradicting the first claim for sufficiently large p. We form two sets of triples as follows. For each a, take b to be any node whose distance from a is in the range $[\ell_a, \frac{1}{2}(\ell_a + u_a)]$, and take c to be any node among the βn farthest nodes from a. Also, for each a, take b to be any node among the nearest βn nodes from a, and take c to be any node whose distance from a is in the range $[\frac{1}{2}(\ell_a + u_a), u_a]$. All of these triples are skew because they span a distance of at least $\frac{1}{2}(u_a - \ell_a) > p\lg n$. The total number of these triples is at least $n(1-2\beta)n\beta n$: n choices for a, $(1-2\beta)n$ choices for either b or c from the middle range $[\ell_a, u_a]$, and βn choices for either c or b from the farthest or nearest βn nodes from a. Therefore $(1-2\beta)\beta n^3$ triples (a,b,c) are skew, which by the first claim is a contradiction provided $p \ge 4/((1-2\beta)\beta)$.

Third we claim that roughly the same property holds for all a'. More precisely, let a be the node from the previous claim, and suppose all nodes b among the middle $(1-2\beta)n$ have distances in an interval $D \pm p\lg n$. We claim that, for every vertex a', if we ignore the nearest $3\beta n$ and farthest $3\beta n$ nodes from a', then the remaining nodes b have distances $d(a,b)$ in the interval $D \pm 2p\lg n$. Otherwise, for some vertex a', either the nearest $3\beta n$ nodes have distances from a' less than $D - 2p\lg n$ or the farthest $3\beta n$ nodes have distances from a' more than $D + 2p\lg n$ (or both). Among these $3\beta n$ nodes, at most $2\beta n$ of them are among the nearest βn or farthest βn nodes from a. Thus, we obtain βn nodes whose distance from a is in the interval $D \pm p\lg n$ but whose distance from a' is outside the interval $D \pm 2p\lg n$. Thus adding the edge aa' decreases the sum of distances from a and a' by at least $\beta np\lg n$ total. By averaging, the sum of distances from either a or a' decreases by at least $\frac{1}{2}\beta np\lg n$. Relabel a and a' so that the sum of distances from a' so improves. By Lemma 10, and because $d > 2\lg n$, there is an edge xy where $d(a',x) \le \lg n$ and whose removal increases the sum of distances from x by at most $2n(1+\lg n)$. Because $d(a',x) \le \lg n$, adding edge xa still improves the sum of distances from x by at least $\beta n(\frac{1}{2}p - 1)\lg n$. Thus, swapping edge xy with edge xa improves the sum of distances from x by at least $\beta n(\frac{1}{2}p - 1)\lg n - 2n(1+\lg n)$. If $p \ge 8/\beta$, this improvement is at least $(2-\beta)n\lg n - 2n$. Because $\beta \le 1$, this improvement is at least $n\lg n - 2n$, which is positive for $n \ge 5$, contradicting equilibrium.

Finally we take the xth power of the graph for an integer x. If two vertices a and b have distance $d(a,b)$ in the original graph, then they have distance $\lceil d(a,b)/x \rceil$ in the power graph. In other words, the power-graph construction coalesces distances between consecutive integer multiples of x down to a common distance (the larger of the two multiples). Therefore, choosing $x = 2p\lg n + 1$ implies that all distances in the range $D \pm 2p\lg n$ convert to at most two distances r and $r+1$. As argued by the previous claim, every vertex a has at least $(1-6\beta)n$ vertices b within this distance range, mapping to distances of either r or $r+1$ in the power graph. The diameter of the power graph is $\lceil d/x \rceil = \Theta(d/(p\lg n)) = \Theta(\beta d/\lg n)$.

To obtain distances of just r, we need a power x with the property that no integer multiple of x falls in the interval $D \pm 2p \lg n$. We show that this is possible for $x = O(\lg^2 n)$: for any interval $I = [i, j]$ where $|j - i| = O(\lg n)$ and $0 < i, j < n$, there is a number $x = O(\lg^2 n)$ such that no integer multiple of x is in $[i, j]$. By a simple consequence of the prime number theorem, the product P of all primes in $[1, y]$ is $e^{(1+o(1))y}$. On the other hand, the product P' of all members of the interval I is at most $n^{|I|} \le e^{O(\lg^2 n)}$. For $y = c \lg^2 n$ with an appropriate constant c, P exceeds P', implying that there is a prime $x \le c \lg^2 n$ that does not divide P' and hence does not divide any member of I, as needed. □

Conjecture 14 *Distance-almost-uniform graphs have diameter $O(\lg n)$.*

If Conjecture 14 is true, Theorem 13 implies that sum equilibrium graphs have diameter $O(\lg^2 n)$. (The slightly weaker conjecture for distance-uniform graphs implies an upper bound of $O(\lg^3 n)$.) Note that for Conjecture 14 it is crucial that we require *every* vertex to have distance exactly r to almost every vertex, not just that almost all pairs of vertices have distance exactly r. Otherwise, a large-diameter example would be a node of degree $\Theta(1/\varepsilon)$ attached to paths of length $(d-2)/2$, with $\Theta(\varepsilon n)$ vertices attached to the end of each path. Provided $d = O(\varepsilon n)$, the number of vertices can be made $\Theta(n)$.

While we have not been able to prove or disprove Conjecture 14, we can prove it (in a strong form) for Cayley graphs of Abelian groups. Recall that the Cayley graph of an Abelian group A with respect to a set $S \subset A$ satisfying $S = -S$ is a graph in which the set of vertices is the set of all elements of the group A, where a, a' are adjacent if and only if there exists an $s \in S$ so that $a + s = a'$. Thus, for example, the graph described in Section 4 is the Cayley graph of the group of all elements of Z_{2k}^2 with an even sum of coordinates, with respect to the generating set $S = \{(1, 1), (1, -1), (-1, 1), (-1, -1)\}$.

Theorem 15 *Let G be an ε-distance-uniform graph with n vertices, and suppose that G is a Cayley graph of an Abelian group and that $\varepsilon < 1/4$. Then the diameter of G is at most $O\left(\frac{\lg n}{\lg(1/\varepsilon)}\right)$.*

Proof: Let G be the Cayley graph of the Abelian group A with respect to the set S. For each integer $i \ge 1$ put

$$iS = \{s_1 + s_2 + \cdots + s_i : s_j \in S \text{ for all } 1 \le j \le i\}.$$

Note that iS is the set of all vertices of G that can be reached from the element $0 \in A$ by a walk of length i. Since G is ε-distance-uniform, there is an integer r so that all vertices of G but at most εn are of distance r or $r + 1$ from 0. Therefore $|(r - 1)S| \le \varepsilon n$, while $|(r + 1)S| \ge (1 - \varepsilon)n$.

A known consequence of the Plunnecke Inequalities (see, e.g., [TV06]), which can also be derived from the results in [Alo03, Section 2], is that if S is a subset of an Abelian group then for every $q > p$, $|qS| \le |pS|^{q/p}$. Applying it in our setting with $q = r + 1$ and $p = r - 1$ we conclude that

$$(1 - \varepsilon)n \le |(r + 1)S| \le |(r - 1)S|^{1+\frac{2}{r-1}} \le \varepsilon n \cdot n^{2/(r-1)}.$$

Therefore $\lg(\frac{1-\varepsilon}{\varepsilon}) \le \frac{2}{r-1} \lg n$, implying that $r \le O\left(\frac{\lg n}{\lg(1/\varepsilon)}\right)$. The desired result follows, as the diameter of G is clearly at most $2r + 2$. □

Acknowledgments

We thank an anonymous referee for helpful comments.

References

[Alb09] Susanne Albers. On the value of coordination in network design. *SIAM Journal on Computing*, 38(6):2273–2302, February 2009.

[AEED+06] Susanne Albers, Stefan Eilts, Eyal Even-Dar, Yishay Mansour, and Liam Roditty. On Nash equilibria for a network creation game. In *Proceedings of the 17th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 89–98, Miami, Florida, 2006.

[Alo03] Noga Alon. Problems and results in extremal combinatorics. I. *Discrete Mathematics*, 273(1-3):31–53, 2003.

[AFM07] Nir Andelman, Michal Feldman, and Yishay Mansour. Strong price of anarchy. In *Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 189–198, 2007.

[CP05] Jacomo Corbo and David Parkes. The price of selfish behavior in bilateral network formation. In *Proceedings of the 24th Annual ACM Symposium on Principles of Distributed Computing*, pages 99–107, Las Vegas, Nevada, 2005.

[Dem02] Erik D. Demaine. Cache-oblivious algorithms and data structures. In *Lecture Notes from the EEF Summer School on Massive Data Sets*, Aarhus, Denmark, June 2002.

[DHMZ09] Erik D. Demaine, MohammadTaghi Hajiaghayi, Hamid Mahini, and Morteza Zadimoghaddam. The price of anarchy in cooperative network creation games. *ACM SIGecom Exchanges*, 8(2), December 2009.

[DHMZ12] Erik D. Demaine, MohammadTaghi Hajiaghayi, Hamid Mahini, and Morteza Zadimoghaddam. The price of anarchy in network creation games. *ACM Transactions on Algorithms*, 8(2):Paper 13, 2012.

[FLM+03] Alex Fabrikant, Ankur Luthra, Elitza Maneva, Christos H. Papadimitriou, and Scott Shenker. On a network creation game. In *Proceedings of the 22nd Annual Symposium on Principles of Distributed Computing*, pages 347–351, Boston, Massachusetts, 2003.

[FLPR99] Matteo Frigo, Charles E. Leiserson, Harald Prokop, and Sridhar Ramachandran. Cache-oblivious algorithms. In *Proceedings of the 40th Annual IEEE Symposium on Foundations of Computer Science*, pages 285–297, New York, October 1999.

[HM07] Yair Halevi and Yishay Mansour. A network creation game with nonuniform interests. In *Proceedings of the 3rd International Workshop on Internet and Network Economics*, volume 4858 of *Lecture Notes in Computer Science*, pages 287–292, San Diego, CA, December 2007.

[LPR+08] Nikolaos Laoutaris, Laura J. Poplawski, Rajmohan Rajaraman, Ravi Sundaram, and Shang-Hua Teng. Bounded budget connection (BBC) games or how to make friends and influence people, on a budget. In *Proceedings of the 27th ACM Symposium on Principles of Distributed Computing*, pages 165–174, 2008.

[TV06] Terence Tao and Van Vu. *Additive Combinatorics*. Cambridge University Press, Cambridge, 2006.